Mixed quantum state detection with inconclusive results

Yonina C. Eldar*
Technion—Israel Institute of Technology, Technion City, Haifa 32000, Israel
(Dated: January 29, 2003)

We consider the problem of designing an optimal quantum detector with a fixed rate of inconclusive results that maximizes the probability of correct detection, when distinguishing between a collection of mixed quantum states. We develop a sufficient condition for the *scaled inverse measurement* to maximize the probability of correct detection for the case in which the rate of inconclusive results exceeds a certain threshold. Using this condition we derive the optimal measurement for linearly independent pure-state sets, and for mixed-state sets with a broad class of symmetries. Specifically, we consider geometrically uniform (GU) state sets and compound geometrically uniform (CGU) state sets with generators that satisfy a certain constraint.

We then show that the optimal measurements corresponding to GU and CGU state sets with arbitrary generators are also GU and CGU respectively, with generators that can be computed very efficiently in polynomial time within any desired accuracy by solving a semidefinite programming problem.

PACS numbers: 03.67.Hk

I. INTRODUCTION

Quantum information theory refers to the distinctive information processing properties of quantum systems, which arise when information is stored in or retrieved from quantum states. A fundamental aspect of quantum information theory is that non-orthogonal quantum states cannot be perfectly distinguished. Therefore, a central problem in quantum mechanics is to design measurements optimized to distinguish between a collection of non-orthogonal quantum states.

We consider a quantum state ensemble consisting of m positive semidefinite Hermitian density operators $\{\rho_i, 1 \leq i \leq m\}$ on an n-dimensional complex Hilbert space $\mathcal{H}$, with prior probabilities $\{p_i > 0, 1 \leq i \leq m\}$. For our *measurement* we consider general positive operator-valued measures [1, 2], consisting of positive semidefinite Hermitian operators that form a resolution of the identity on $\mathcal{H}$.

Different approaches to distinguishing between the density operators ρ_i have emerged. In one approach, the measurement consists of m measurement operators which are designed to maximize the probability of correct detection. Necessary and sufficient conditions for an optimum measurement maximizing the probability of correct detection have been developed [3, 4, 5]. Closed-form analytical expressions for the optimal measurement have been derived for several special cases [6, 7, 8, 9, 10, 11]. In particular, the optimal measurement for pure and mixed-state ensembles with broad symmetry properties, referred to as geometrically uniform (GU) and compound GU (CGU) state sets, are considered in [7, 11]. Iterative procedures maximizing the probability of correct detection have also been developed for cases in which the optimal measurement cannot be found explicitly [3, 12].

More recently, a different approach to the problem has emerged, which in some cases may be more useful. This approach, referred to as unambiguous quantum state discrimination, combines error free discrimination with a certain fraction of inconclusive results [13, 14, 15, 16, 17, 18, 19, 20]. The basic idea, pioneered by Ivanovic [15], is to design a measurement that with probability β returns an inconclusive result, but such that if the measurement returns an answer, then the answer is correct with probability 1. In this case the measurement consists of $m+1$ measurement operators corresponding to $m+1$ outcomes, where m outcomes correspond to detection of each of the states and the additional outcome corresponds to an inconclusive result. Chefles [13] showed that a necessary and sufficient condition for the existence of unambiguous measurements for distinguishing between a collection of quantum states is that the states are linearly independent pure states. The optimal measurement minimizing the probability β of an inconclusive result when distinguishing between GU and CGU pure-state sets was considered in [14], and was shown under certain conditions to be equal to the equal-probability measurement (EPM).

An interesting alternative approach for distinguishing between a collection of quantum states, first considered by Chefles and Barnet [21] and Zhang *et al.* [22] for pure-state ensembles, and then later extended by Fiurášek and Ježek [23] to mixed-state ensembles, is to allow for a certain probability of an inconclusive result, and then maximize the probability of correct detection. Thus, in this approach, the measurement again consists of $m+1$ measurement outcomes; however, now the outcomes do not necessarily correspond to perfect detection of each of the states. Indeed, if the quantum states are mixed states or linearly dependent pure states, then perfect detection of each of the states is not possible [13]. Nonetheless, by allowing for inconclusive results, a higher probability

*Electronic address: yonina@ee.technion.ac.il

of correct detection can be obtained in comparison with the probability of correct detection attainable without inconclusive results.

Necessary conditions as well as a set of sufficient conditions on the optimal measurement operators maximizing the probability of correct detection subject to the constraint that the probability of an inconclusive result is equal to a constant β were derived in [23], using Lagrange multiplier theory. It was also pointed out in [23] that obtaining a closed form analytical solution to the optimal measurement operators directly form these conditions is a difficult problem.

In this paper we extend the results of [23] in several ways. First, using principles of duality in vector space optimization, in Section III we show that the conditions derived in [23] are both necessary and sufficient. We also show that the Lagrange multipliers can be obtained by solving a reduced size semidefinite programming problem. This approach lends itself to efficient computational methods which are guaranteed to converge to the global optimum.

Second, we derive a general condition in Section IV under which the *scaled inverse measurement (SIM)* is optimal. This measurement consists of measurement operators that are proportional to the reciprocal states associated with the given state ensemble, and can be regarded as a generalization of the EPM to mixed-state ensembles.

Third, we develop the optimal measurement for state sets with broad symmetry properties. Specifically, in Section V we consider GU state sets defined over a finite group of unitary matrices. We obtain a convenient characterization of the SIM and show that the SIM operators have the same symmetries as the original state set. We then show that for a pure GU state set and for values of β exceeding a certain threshold, the SIM is optimal. For a mixed GU state set, under a certain constraint on the generator and for values of β exceeding a threshold, the SIM is again shown to be optimal. For arbitrary values of β, the optimal measurement operators corresponding to a pure or mixed GU state set are shown to be GU with the same generating group, and can be computed very efficiently in polynomial time.

In Section VI we consider CGU state sets [24], in which the states are generated by a group of unitary matrices using *multiple* generators. We obtain a convenient characterization of the SIM for CGU state sets, and show that the SIM vectors are themselves CGU. Under a certain condition on the generators and for values of β exceeding a threshold, the SIM is shown to be optimal. Finally we show that for arbitrary CGU state sets and for arbitrary values of β, the optimal measurement operators are also CGU, and we propose an efficient algorithm for computing the optimal generators.

It is interesting to note that a closed form analytical expression exists for the optimal measurement when distinguishing between GU and CGU (possibly mixed) state sets with generators that satisfy a certain constraint, under each of the three approaches outlined to quantum detection, where in the last approach we assume that β exceeds a certain threshold. Furthermore, as shown in [7, 14] and in Sections V and VI, the optimal measurement operators corresponding to GU and CGU state sets are also GU and CGU respectively, under each one of the three outlined optimality criteria.

Before proceeding to the detailed development, we provide in the next section a statement of our problem.

II. PROBLEM FORMULATION

Assume that a quantum channel is prepared in a quantum state drawn from a collection of given states represented by density operators $\{\rho_i, 1 \leq i \leq m\}$ on an n-dimensional complex Hilbert space $\mathcal{H}$. We assume without loss of generality that the eigenvectors of $\rho_i, 1 \leq i \leq m$, collectively span[37] $\mathcal{H}$ so that $m \geq n$. Since ρ_i is Hermitian and positive semidefinite, we can express ρ_i as $\rho_i = \phi_i \phi_i^*$ for some matrix ϕ_i, *e.g.,* via the Cholesky or eigendecomposition of ρ_i [25]. We refer to ϕ_i as a *factor* of ρ_i. The choice of ϕ_i is not unique; if ϕ_i is a factor of ρ_i, then any matrix of the form $\phi_i' = \phi_i Q_i$ where Q_i is an arbitrary matrix satisfying $Q_i Q_i^* = I$, is also a factor of ρ_i.

To detect the state of the system a measurement is constructed comprising $m+1$ measurement operators $\{\Pi_i, 0 \leq i \leq m\}$ that satisfy

$$\begin{aligned} \Pi_i &\geq 0, \quad 0 \leq i \leq m; \\ \sum_{i=0}^{m} \Pi_i &= I. \end{aligned} \tag{1}$$

Each of the operators $\Pi_i, 1 \leq i \leq m$ correspond to detection of the corresponding states $\rho_i, 1 \leq i \leq m$, and Π_0 corresponds to an inconclusive result. We seek the measurement operators Π_i that maximize the probability of correct detection, subject to the constraint that the probability of an inconclusive result is equal to a constant $\beta < 1$.

Given that the transmitted state is ρ_j, the probability of correctly detecting the state using measurement operators $\{\Pi_i, 1 \leq i \leq m\}$ is $\mathrm{Tr}(\rho_j \Pi_j)$ and the probability of a detection error is $\sum_{i=1, i\neq j}^{m} \mathrm{Tr}(\rho_j \Pi_i)$. Therefore, the probability of correct detection is given by

$$P_D = \sum_{i=1}^{m} p_i \mathrm{Tr}(\rho_i \Pi_i), \tag{2}$$

where $p_i > 0$ is the prior probability of ρ_i, with $\sum_i p_i = 1$, and the probability of a detection error is given by

$$P_E = \sum_{i=1}^{m} \sum_{j=1, j\neq i}^{m} p_i \mathrm{Tr}(\rho_i \Pi_j). \tag{3}$$

The probability of an inconclusive result is

$$P_I = \sum_{i=1}^{m} p_i \mathrm{Tr}(\rho_i \Pi_0) = \mathrm{Tr}(\Delta \Pi_0) = \beta, \tag{4}$$

where for brevity we denote

$$\Delta = \sum_{i=1}^{m} p_i \rho_i. \tag{5}$$

Our problem is to find the measurement operators $\{\Pi_i, 0 \leq i \leq m\}$ that maximize P_D of (2) subject to the constraints (1) and (4).

Note that since $\mathrm{Tr}(\rho_i) = 1$ for all i,

$$P_D + P_E + P_I = \sum_{i=1}^{m} p_i \mathrm{Tr}(\rho_i) = 1. \tag{6}$$

When $P_E = 0$ the states are distinguished unambiguously so that if outcome i is obtained for some $1 \leq i \leq m$, then the state is ρ_i with probability one. It was shown in [13] that with $\beta < 1$ we can choose measurement operators such that $P_E = 0$ if and only if the state ensemble is a linearly independent pure-state ensemble consisting of density operators ρ_i of the form $\rho_i = |\phi_i\rangle\langle\phi_i|$ for a set of linearly independent vectors $|\phi_i\rangle$. If the vectors $|\phi_i\rangle$ are linearly dependent, or if the ensemble is a mixed-state ensemble, then P_E cannot be equal to 0. Nonetheless, we may seek the measurement operators that minimize P_E, or equivalently, maximize P_D, subject to $P_I = \beta$ for some $\beta < 1$.

Equipped with the standard operations of addition and multiplication by real numbers, the space $\mathcal{B}$ of all Hermitian $n \times n$ matrices is an n^2-dimensional *real* vector space. As noted in [23], by choosing an appropriate basis for $\mathcal{B}$, the problem of maximizing P_D subject to (1) and (4) can be put in the form of a standard semidefinite programming problem, which is a convex optimization problem; for a detailed treatment of semidefinite programming problems see, *e.g.,* [26, 27, 28, 29]. Recently, methods based on semidefinite programming have been employed in a variety of different problems in quantum detection and quantum information [3, 14, 30, 31, 32, 33]. By exploiting the many well known algorithms for solving semidefinite programs [29], *e.g.,* interior point methods[38] [26, 28], the optimal measurement can be computed very efficiently in polynomial time.

The semidefinite programming formulation can also be used to derive necessary and sufficient conditions for optimality, which we discuss in the next section.

III. CONDITIONS FOR OPTIMALITY

Using Lagrange multipliers, it was shown in [23] that a set of measurement operators $\{\widehat{\Pi}_i, 0 \leq i \leq m\}$ maximizes P_D subject to $P_I = \beta$ for a state set $\{\rho_i, 1 \leq i \leq m\}$ with prior probabilities $\{p_i, 1 \leq i \leq m\}$ if there exists an Hermitian $\widehat{X}$ and a constant $\hat{\delta}$ satisfying

$$\widehat{X} \geq p_i \rho_i, \quad 1 \leq i \leq m; \tag{7}$$

$$\widehat{X} \geq \hat{\delta} \Delta, \tag{8}$$

such that

$$(\widehat{X} - p_i \rho_i)\widehat{\Pi}_i = 0, \quad 1 \leq i \leq m; \tag{9}$$

$$(\widehat{X} - \hat{\delta}\Delta)\widehat{\Pi}_0 = 0. \tag{10}$$

It was also shown that (9) and (10) are necessary conditions for optimality.

In Appendix A we use duality arguments similar to those used in [3] to show that (7)–(10) are *necessary and sufficient* conditions for optimality, so that a set of measurement operators $\widehat{\Pi}_i$ maximizes P_D subject to $P_I = \beta$ if and only if there exists an Hermitian $\widehat{X}$ and a constant $\hat{\delta}$ satisfying (7)–(10). Furthermore, we show that $\widehat{X}$ and $\hat{\delta}$ can be determined as the solution to the following semidefinite programming problem:

$$\min_{X \in \mathcal{B}, \delta \in \mathcal{R}} tr(X) - \delta\beta, \tag{11}$$

where $\mathcal{R}$ denotes the reals, subject to

$$X \geq p_i \rho_i, \quad 1 \leq i \leq m;$$
$$X \geq \delta \Delta. \tag{12}$$

The problem of (11)–(12) is referred to as the dual problem.

Note that the dual problem involves many fewer decision variables than the primal maximization problem. Specifically, in the dual problem we have $n^2 + 1$ real decision variables while the primal problem has $(m+1)n^2$ real decision variables. Therefore, it is advantageous to solve the dual problem and then use (9) and (10) to determine the optimal measurement operators, rather than solving the primal problem directly.

The necessary conditions (7) and (9) together imply that the rank of each optimal measurement operator is no larger than the rank of the corresponding density operator; see [3]. In particular, if the quantum state ensemble is a pure-state ensemble consisting of (not necessarily independent) rank-one density operators $\rho_i = |\phi_i\rangle\langle\phi_i|$, then the optimal measurement is a pure-state measurement consisting of rank-one measurement operators $\Pi_i = |\mu_i\rangle\langle\mu_i|$.

As pointed out in [23], obtaining a closed-form analytical expression for the optimal measurement operators directly from the necessary and sufficient conditions for optimality is a difficult problem. Since (11) is a (convex) semidefinite programming [26, 28, 29] problem, there are very efficient methods for solving (11). In particular, the optimal matrix $\widehat{X}$ and optimal scalar $\hat{\delta}$ minimizing $\mathrm{Tr}(X) - \delta\beta$ subject to (12) can be computed in Matlab using the linear matrix inequality (LMI) Toolbox (see [3, 14] for further details). Once we determine $\widehat{X}$, the optimal measurement operators $\widehat{\Pi}_i$ can be computed in a similar manner to that described in [3].

A suboptimal measurement that has been suggested as a detection measurement for unambiguous quantum state discrimination between linearly independent pure

quantum states, is the EPM [13, 14, 20], in which the measurement vectors are proportional to the reciprocal states associated with the states to be distinguished. A general condition under which the EPM is optimal for distinguishing between pure quantum states was derived in [14]. It was also shown that for GU state sets and for CGU state sets with generators satisfying a certain constraint, the EPM is optimal.

In the next section we consider a generalization of the EPM to mixed quantum states, which we refer to as the *scaled inverse measurement (SIM)*. We then use the necessary and sufficient conditions for optimality to derive a general condition under which the SIM is optimal. In Sections V and VI we consider some special cases of mixed and pure-state sets for which the SIM is optimal, and derive explicit formulas for the optimal measurement operators.

IV. THE SIM AND THE OPTIMAL MEASUREMENT

The SIM corresponding to a set of density operators $\{\rho_i = \phi_i\phi_i^*, 1 \le i \le m\}$ with eigenvectors that collectively span $\mathcal{H}$ and prior probabilities $\{p_i, 1 \le i \le m\}$ consists of the measurement operators $\{\Sigma_i = \mu_i\mu_i^*, 0 \le i \le m\}$ where

$$\mu_i = \gamma(\Psi\Psi^*)^{-1}\psi_i = \gamma\Delta^{-1}\psi_i, \quad 1 \le i \le m, \tag{13}$$

for some $\gamma > 0$ and $\Sigma_0 = I - \sum_{i=1}^m \mu_i\mu_i^*$. Here Ψ is the matrix of (block) columns $\psi_i = \sqrt{p_i}\phi_i$. Note that since the eigenvectors of the $\{\rho_i\}$ collectively span $\mathcal{H}$, the columns of the $\{\psi_i\}$ also together span $\mathcal{H}$, so $\Psi\Psi^*$ is invertible. From (13),

$$\sum_{i=1}^m \mu_i\mu_i^* = \gamma^2\Delta^{-1}\left(\sum_{i=1}^m \psi_i\psi_i^*\right)\Delta^{-1} = \gamma^2\Delta^{-1}, \tag{14}$$

so that

$$\Sigma_0 = I - \sum_{i=1}^m \mu_i\mu_i^* = I - \gamma^2\Delta^{-1}. \tag{15}$$

If follows from (15) that the SIM operators satisfy (1) if and only if $\gamma^2 \le \lambda_n$ where $\{\lambda_i, 1 \le i \le n\}$ denote the eigenvalues of $\Delta = \Psi\Psi^*$ and $\lambda_n = \min\lambda_i$.

In the case in which the prior probabilities are all equal,

$$\mu_i = \gamma(\Phi\Phi^*)^{-1}\phi_i, \quad 1 \le i \le m, \tag{16}$$

where Φ is the matrix of (block) columns ϕ_i.

Since the factors ϕ_i are not unique, the SIM factors μ_i are also not unique. If μ_i are the SIM factors corresponding to ϕ_i, then the SIM factors corresponding to $\phi_i' = \phi_i Q_i$ with $Q_iQ_i^* = I$ are $\mu_i' = \mu_i Q_i$. Therefore, although the SIM factors are not unique, the SIM operators $\Sigma_i = \mu_i\mu_i^*$ are unique.

The SIM corresponding to a pure-state ensemble $|\phi_i\rangle$ consists of the measurement vectors $|\mu_i\rangle = \gamma\Delta^{-1}|\psi_i\rangle$, where $|\psi_i\rangle = \sqrt{p_i}|\phi_i\rangle$. If in addition $\gamma = \sqrt{\lambda_n}$, then the SIM vectors are equal to the EPM vectors [14].

The probability of an inconclusive result using the SIM is

$$P_I = \mathrm{Tr}(\Delta\Sigma_0) = \mathrm{Tr}(\Delta) - \gamma^2\mathrm{Tr}(I) = 1 - n\gamma^2. \tag{17}$$

Therefore to satisfy (4),

$$\gamma = \sqrt{\frac{1-\beta}{n}}. \tag{18}$$

Since we must also have $\gamma^2 \le \lambda_n$ we conclude that β must satisfy

$$\beta \ge 1 - n\lambda_n \triangleq \beta_{\min}. \tag{19}$$

For linearly independent pure quantum states it was shown in [14] that the SIM with $\gamma = \sqrt{\lambda_n}$ minimizes the probability of an inconclusive result subject to the constraint that $P_D = 1$ for state sets with strong symmetry properties. The smallest possible probability of an inconclusive result in this case is $\beta = \beta_{\min}$. It turns out that for a large class of state sets, including those discussed in [14], the SIM also maximizes P_D subject to $P_I = \beta$ for $\beta \ge \beta_{\min}$. From the necessary and sufficient conditions for optimality discussed in Section III it follows that the SIM is optimal if and only if the measurement operators $\widehat{\Pi}_i = \mu_i\mu_i^*, 1 \le i \le m$ and $\widehat{\Pi}_0 = \Sigma_0$ defined by (13) and (15) satisfy (9) and (10) for some Hermitian $\widehat{X}$ and constant $\hat{\delta}$ satisfying (7) and (8). A sufficient condition for optimality of the SIM is given in the following theorem, the proof of which is provided in the Appendix.

Theorem 1. *Let $\{\rho_i = \phi_i\phi_i^*, 1 \le i \le m\}$ denote a collection of quantum states with prior probabilities $\{p_i, 1 \le i \le m\}$. Let $\{\Sigma_i = \mu_i\mu_i^*, 0 \le i \le m\}$ with $\{\mu_i = \gamma\Delta^{-1}\psi_i, 1 \le i \le m\}$ and $\Sigma_0 = I - \sum_{i=1}^m \Sigma_i$ denote the scaled inverse measurement (SIM) operators corresponding to $\{\psi_i = \sqrt{p_i}\phi_i, 1 \le i \le m\}$, where $\gamma^2 = (1-\beta)/n$, $\Delta = \Psi\Psi^*$ and Ψ is the matrix with block columns ψ_i. Let $\lambda_n = \min\lambda_i$ where λ_i are the eigenvalues of Δ. Then the SIM maximizes P_D subject to $P_I = \beta$ for $\beta \ge \beta_{\min}$ with $\beta_{\min} = 1 - n\lambda_n$ if for each $1 \le i \le m$, $(1/\gamma)\mu_i^*\psi_i = \psi_i^*\Delta^{-1}\psi_i = \alpha I$, where α is a constant independent of i.*

It is interesting to note that the condition of Theorem 1 is identical to the condition given in Theorem 1 of [7] for the least-squares measurement, or the square-root measurement, to maximize the probability of correct detection when $P_I = 0$.

As we expect, the condition $\psi_i^*\Delta^{-1}\psi_i = \alpha I$ does not depend on the choice of factor ϕ_i. Indeed, if $\phi_i' = \phi_i Q_i$ is another factor of ρ_i with Q_i satisfying $Q_iQ_i^* = I$, and if Ψ' is the matrix of block columns $\psi_i' = \sqrt{p_i}\phi_i' = \sqrt{p_i}\phi_i Q_i$, then it is easy to see that $(\psi_i')^*(\Psi'\Psi'^*)^{-1}\psi_i' = \alpha I$ if and only if $\psi_i^*\Delta^{-1}\psi_i = \alpha I$.

For a pure-state ensemble consisting of density operators $\rho_i = |\phi_i\rangle\langle\phi_i|$ for a set of vectors $|\phi_i\rangle$, $\langle\psi_i|\Delta^{-1}|\psi_i\rangle$ is the ith diagonal element of $P = \Psi^*\Delta^{-1}\Psi = \Psi^*(\Psi\Psi^*)^{-1}\Psi$. The matrix P is just the orthogonal projection onto $\mathcal{N}(\Psi)^\perp$, where $\mathcal{N}(\Psi)$ is the null space of Ψ. If the vectors $|\phi_i\rangle$ are linearly independent, then $\mathcal{N}(\Psi)^\perp = \{0\}$ so that $P = I$ and $\langle\psi_i|\Delta^{-1}|\psi_i\rangle = 1$ for all i. It therefore follows from Theorem 1 that for a pure-state ensemble consisting of linearly independent state vectors, the SIM maximizes P_D subject to $P_I = \beta$ for any $\beta \geq \beta_{\min}$.

If the state $\rho_i = \phi_i\phi_i^*$ is transmitted with prior probability p_i, then the probability of correctly detecting the state using measurement operators $\Sigma_i = \mu_i\mu_i^*$ is $p_i\mathrm{Tr}(\mu_i^*\phi_i\phi_i^*\mu_i) = \mathrm{Tr}(\mu_i^*\psi_i\psi_i^*\mu_i)$. It follows that if the condition for optimality of Theorem 1 is met, then the probability of correctly detecting each of the states ρ_i using the SIM is the same.

For a pure-state ensemble consisting of states $|\phi_i\rangle$ with prior probabilities p_i, the probability of correct detection of the ith state is given by $|\langle\mu_i|\psi_i\rangle|^2$. Since $\langle\mu_i|\psi_i\rangle = \gamma\langle\psi_i|\Delta^{-1}|\psi_i\rangle \geq 0$ for any set of weighted vectors $|\psi_i\rangle$, $\langle\mu_i|\psi_i\rangle$ is constant for all i if and only if $|\langle\mu_i|\psi_i\rangle|^2$ is constant for all i. Therefore, we may interpret the condition in Theorem 1 for pure-state ensembles as follows: The SIM is optimal for a set of states $|\phi_i\rangle$ with prior probabilities p_i and for $\beta \geq \beta_{\min}$ if the probability of detecting each one of the states using the SIM vectors is the same, regardless of the specific state chosen.

In the remainder of the paper we use Theorem 1 to derive the optimal measurement for mixed and (not necessarily independent) pure-state sets with certain symmetry properties. The symmetry properties we consider are quite general, and include many cases of practical interest.

V. GEOMETRICALLY UNIFORM STATE SETS

In this section we consider *geometrically uniform (GU)* [34] state sets in which the density operators ρ_i are defined over a group of unitary matrices and are generated by a single generating matrix. We first obtain a convenient characterization of the SIM for GU state sets, and show that under a certain constraint on the generator the SIM is optimal when $\beta \geq \beta_{\min}$. In particular, for (not necessarily independent) pure-state ensembles the SIM is optimal. We then show that for arbitrary GU state sets and arbitrary values of β, the optimal measurement is also GU, and we develop an efficient computational method for finding the optimal generators.

Let $\mathcal{G} = \{U_i, 1 \leq i \leq m\}$ be a finite group of m unitary matrices U_i. That is, $\mathcal{G}$ contains the identity matrix I; if $\mathcal{G}$ contains U_i, then it also contains its inverse $U_i^{-1} = U_i^*$; and the product U_iU_j of any two elements of $\mathcal{G}$ is in $\mathcal{G}$ [35].

A state set generated by $\mathcal{G}$ using a single generating operator ρ is a set $\mathcal{S} = \{\rho_i = U_i\rho U_i^*, U_i \in \mathcal{G}\}$. The group $\mathcal{G}$ is the *generating group* of $\mathcal{S}$. Such a state set has strong symmetry properties and is called GU. For consistency with the symmetry of $\mathcal{S}$, we will assume equiprobable prior probabilities on $\mathcal{S}$.

If the state set $\{\rho_i, 1 \leq i \leq m\}$ is GU, then we can always choose factors ϕ_i of ρ_i such that $\{\phi_i = U_i\phi, U_i \in \mathcal{G}\}$ where ϕ is a factor of ρ, so that the factors ϕ_i are also GU with generator ϕ. In the remainder of this section we explicitly assume that the factors are chosen to be GU.

A. Optimality of the SIM for GU States

For a GU state set with generating group $\mathcal{G}$, $\Phi\Phi^*$ commutes with each of the matrices $U_i \in \mathcal{G}$ [7, 24]. Consequently, $T = (\Phi\Phi^*)^{-1}$ also commutes with U_i for all i, so that from (16)

$$\mu_i = \gamma T\phi_i = \gamma TU_i\phi = \gamma U_iT\phi = U_i\mu,\ 1 \leq i \leq m, \quad (20)$$

where

$$\mu = \gamma(\Phi\Phi^*)^{-1}\phi. \quad (21)$$

It follows that the SIM factors μ_i are also GU with generating group $\mathcal{G}$ and generator μ given by (21). Therefore, to compute the SIM factors for a GU state set all we need is to compute the generator μ. The remaining measurement factors are then obtained by applying the group $\mathcal{G}$ to μ.

From (20) we have that

$$(1/\gamma)\mu_i^*\psi_i = \frac{1}{\gamma\sqrt{m}}\mu^*U_i^*U_i\phi = \frac{1}{\gamma\sqrt{m}}\mu^*\phi, \quad (22)$$

where ϕ and μ are the generators of the state factors and the SIM factors, respectively. Thus, the probability of correct detection of each one of the states ρ_i using the SIM is the same, regardless of the state transmitted. This then implies from Theorem 1 that for a (not necessarily independent) pure-state GU ensemble the SIM is optimal when $\beta \geq \beta_{\min}$. For a mixed-state ensemble, if the generator ϕ satisfies

$$\phi^*(\Phi\Phi^*)^{-1}\phi = \alpha I \quad (23)$$

for some α, then from Theorem 1 the SIM is again optimal.

B. Optimal Measurement for Arbitrary GU States

If the generator ϕ does not satisfy (23), or if $\beta < \beta_{\min}$, then the SIM is no longer guaranteed to be optimal. Nonetheless, as we now show, the optimal measurement operators that maximize P_D subject to $P_I = \beta$ for any β are GU with generating group $\mathcal{G}$. The corresponding generator can be computed very efficiently in polynomial time.

Suppose that the optimal measurement operators that maximize

$$J(\{\Pi_i\}) = \sum_{i=1}^{m} \mathrm{Tr}(\rho_i \Pi_i), \tag{24}$$

subject to

$$P_I(\{\Pi_i\}) = 1 - \frac{1}{m}\mathrm{Tr}\left(\sum_{i,j=1}^{m} \rho_i \Pi_j\right) = \beta, \tag{25}$$

are $\widehat{\Pi}_i$ and let $\widehat{J} = J(\{\widehat{\Pi}_i\})$. Let $r(j,i)$ be the mapping from $\mathcal{I}\times\mathcal{I}$ to $\mathcal{I}$ with $\mathcal{I} = \{1,\ldots,m\}$, defined by $r(j,i) = k$ if $U_j^* U_i = U_k$. Then the measurement operators $\widehat{\Pi}_i^{(j)} = U_j \widehat{\Pi}_{r(j,i)} U_j^*, 1 \le i \le m$ and $\widehat{\Pi}_0^{(j)} = I - \sum_{i=1}^m \widehat{\Pi}_i^{(j)}$ for any $1 \le j \le m$ are also optimal. Indeed, since $\widehat{\Pi}_i \ge 0, 1 \le i \le m$ and $\sum_{i=1}^m \widehat{\Pi}_i \le I$, $\widehat{\Pi}_i^{(j)} \ge 0, 1 \le i \le m$ and

$$\sum_{i=1}^{m} \widehat{\Pi}_i^{(j)} = U_j \left(\sum_{i=1}^{m} \widehat{\Pi}_i\right) U_j^* \le U_j U_j^* = I. \tag{26}$$

Using the fact that $\rho_i = U_i \rho U_i^*$ for some generator ρ,

$$\begin{aligned} J(\{\widehat{\Pi}_i^{(j)}\}) &= \sum_{i=1}^{m} \mathrm{Tr}(\rho U_i^* U_j \widehat{\Pi}_{r(j,i)} U_j^* U_i) \\ &= \sum_{k=1}^{m} \mathrm{Tr}(\rho U_k^* \widehat{\Pi}_k U_k) \\ &= \sum_{i=1}^{m} \mathrm{Tr}(\rho_i \widehat{\Pi}_i) \\ &= \widehat{J}. \end{aligned} \tag{27}$$

Finally,

$$\begin{aligned} \mathrm{Tr}\left(\sum_{i,s=1}^{m} \rho_i \widehat{\Pi}_s^{(j)}\right) &= \mathrm{Tr}\left(\sum_{i,s=1}^{m} U_j^* U_i \rho U_i^* U_j \widehat{\Pi}_{r(j,s)}\right) \\ &= \mathrm{Tr}\left(\sum_{i,k=1}^{m} U_i \rho U_i^* \widehat{\Pi}_k\right) \\ &= \mathrm{Tr}\left(\sum_{i,k=1}^{m} \rho_i \widehat{\Pi}_k\right), \end{aligned} \tag{28}$$

so that from (25), $P_I(\{\widehat{\Pi}_i^{(j)}\}) = P_I(\{\widehat{\Pi}_i\})$.

Since the measurement operators $\widehat{\Pi}_i^{(j)}$ are optimal for any j, it follows immediately that the measurement operators $\{\overline{\Pi}_i = (1/m)\sum_{j=1}^m \widehat{\Pi}_i^{(j)}, 1 \le i \le m\}$ and $\overline{\Pi}_0 = I - \sum_{i=1}^m \overline{\Pi}_i$ are also optimal. Now, for any $1 \le i \le m$,

$$\begin{aligned} \overline{\Pi}_i &= \frac{1}{m}\sum_{j=1}^{m} U_j \widehat{\Pi}_{r(j,i)} U_j^* \\ &= \frac{1}{m}\sum_{k=1}^{m} U_i U_k^* \widehat{\Pi}_k U_k U_i^* \\ &= U_i \left(\frac{1}{m}\sum_{k=1}^{m} U_k^* \widehat{\Pi}_k U_k\right) U_i^* \\ &= U_i \widehat{\Pi} U_i^*, \end{aligned} \tag{29}$$

where $\widehat{\Pi} = (1/m)\sum_{k=1}^m U_k^* \widehat{\Pi}_k U_k$.

We therefore conclude that the optimal measurement operators can always be chosen to be GU with the same generating group $\mathcal{G}$ as the original state set. Thus, to find the optimal measurement operators all we need is to find the optimal generator $\widehat{\Pi}$. The remaining operators are obtained by applying the group $\mathcal{G}$ to $\widehat{\Pi}$.

Since the optimal measurement operators satisfy $\Pi_i = U_i \Pi U_i^*, 1 \le i \le m$ and $\rho_i = U_i \rho U_i^*$, $\mathrm{Tr}(\rho_i \Pi_i) = \mathrm{Tr}(\rho\Pi)$, so that the problem (2) reduces to the maximization problem

$$\max_{\Pi \in \mathcal{B}} \mathrm{Tr}(\rho\Pi), \tag{30}$$

where $\mathcal{B}$ is the set of $n \times n$ Hermitian operators, subject to the constraints

$$\begin{aligned} &\Pi \ge 0; \\ &\sum_{i=1}^{m} U_i \Pi U_i^* \le I; \\ &1 - \mathrm{Tr}\left(\sum_{i=1}^{m} U_i \rho U_i \Pi\right) = \beta. \end{aligned} \tag{31}$$

The problem of (30) and (31) is a (convex) semidefinite programming problem, and therefore the optimal Π can be computed very efficiently in polynomial time within any desired accuracy [26, 28, 29], for example using the LMI toolbox on Matlab. Note that the problem of (30) and (31) has n^2 real unknowns and 3 constraints, in contrast with the original maximization problem (2) subject to (1) and (4) which has mn^2 real unknowns and $m+2$ constraints.

We summarize our results regarding GU state sets in the following theorem:

Theorem 2 (GU state sets). *Let $\mathcal{S} = \{\rho_i = U_i \rho U_i^*, U_i \in \mathcal{G}\}$ be a geometrically uniform (GU) state set on an n-dimensional Hilbert space, generated by a finite group $\mathcal{G}$ of unitary matrices, where $\rho = \phi\phi^*$ is an arbitrary generator, and let Φ be the matrix of columns $\phi_i = U_i \phi$. Then the scaled inverse measurement (SIM) is given by the measurement operators $\Sigma_i = \mu_i \mu_i^*, 0 \le i \le m$ with*

$$\mu_i = U_i \mu, \quad 1 \le i \le m,$$

where

$$\mu = \gamma(\Phi\Phi^*)^{-1}\phi,$$

with $\gamma^2 = (1-\beta)/n$, *and* $\Sigma_0 = I - \sum_{i=1}^m \mu_i\mu_i^*$. *The SIM has the following properties:*

1. *The measurement operators* $\Sigma_i, 1 \le i \le m$ *are GU with generating group* $\mathcal{G}$;

2. *The probability of correctly detecting each of the states* ρ_i *using the SIM is the same;*

3. *If* $\phi^*(\Phi\Phi^*)^{-1}\phi = \alpha I$ *for some* α, *then the SIM maximizes* P_D *subject to* $P_I = \beta$ *for* $\beta \ge 1 - n\lambda_n$ *where* λ_n *is the smallest eigenvalue of* $(1/m)\sum_{i=1}^m \rho_i$; *In particular, if* $\phi = |\phi\rangle$ *is a vector so that the state set is a pure-state ensemble, then the SIM maximizes* P_D *subject to* $P_I = \beta$ *for any* $\beta \ge 1 - n\lambda_n$.

For an arbitrary generator ϕ *the optimal measurement operators* $\widehat{\Pi}_i, 1 \le i \le m$ *that maximize* P_D *subject to* $P_I = \beta$ *for any* β *are also GU with generating group* $\mathcal{G}$ *and generator* Π *that maximizes* $Tr(\rho\Pi)$ *subject to* $\Pi \ge 0, \sum_{i=1}^m U_i\Pi U_i^* \le I$, *and* $Tr(\sum_{i=1}^m U_i\rho U_i\Pi) = 1 - \beta$.

VI. COMPOUND GEOMETRICALLY UNIFORM STATE SETS

We now consider *compound geometrically uniform (CGU)* [24] state sets which consist of subsets that are GU. As we show, the SIM operators are also CGU so that they can be computed using a *set* of generators. Under a certain condition on the generators and for $\beta \ge \beta_{\min}$, we show that the optimal measurement associated with a CGU state set is equal to the SIM. For arbitrary CGU state sets and arbitrary values of β we show that the optimal measurement operators are CGU, and we derive an efficient computational method for finding the optimal generators.

A CGU state set is defined as a set of density operators $\mathcal{S} = \{\rho_{ik} = \phi_{ik}\phi_{ik}^*, 1 \le i \le l, 1 \le k \le r\}$ such that $\rho_{ik} = U_i\rho_k U_i^*$, where the matrices $\{U_i, 1 \le i \le l\}$ are unitary and form a group $\mathcal{G}$, and the operators $\{\rho_k, 1 \le k \le r\}$ are the generators. We assume equiprobable prior probabilities on $\mathcal{S}$.

If the state set $\{\rho_{ik}, 1 \le i \le l, 1 \le k \le r\}$ is CGU, then we can always choose factors ϕ_{ik} of ρ_{ik} such that $\{\phi_{ik} = U_i\phi_k, 1 \le i \le l\}$ where ϕ_k is a factor of ρ_k, so that the factors ϕ_{ik} are also CGU with generators $\{\phi_k, 1 \le k \le r\}$. In the remainder of this section we explicitly assume that the factors are chosen to be CGU.

A CGU state set is in general not GU. However, for every k, the matrices $\{\phi_{ik}, 1 \le i \le l\}$ and the operators $\{\rho_{ik}, 1 \le i \le l\}$ are GU with generating group $\mathcal{G}$. Examples of CGU state sets are considered in [7].

A. Optimality of the SIM for CGU State Sets

With Φ denoting the matrix of (block) columns ϕ_{ik}, it was shown in [7, 24] that $\Phi\Phi^*$, and consequently $T = (\Phi\Phi^*)^{-1}$, commutes with each of the matrices $U_i \in \mathcal{G}$. Thus, the SIM operators are $\Sigma_{ik} = \mu_{ik}\mu_{ik}^*, 1 \le i \le l, 1 \le k \le r$ with

$$\mu_{ik} = \gamma T\phi_{ik} = \gamma TU_i\phi_k = U_i\mu_k, \tag{32}$$

where

$$\mu_k = \gamma T\phi_k = \gamma(\Phi\Phi^*)^{-1}\phi_k. \tag{33}$$

Therefore the SIM factors are also CGU with generating group $\mathcal{G}$ and generators μ_k given by (33). To compute the SIM factors all we need is to compute the generators μ_k. The remaining measurement factors are then obtained by applying the group $\mathcal{G}$ to each of the generators.

From (32),

$$\mu_{ik}^*\phi_{ik} = \mu_k^*U_i^*U_i\phi_k = \mu_k^*\phi_k, \tag{34}$$

so that from Theorem 1 the SIM is optimal if

$$\mu_k^*\phi_k = \gamma\phi_k^*T\phi_k = \alpha I, \quad 1 \le k \le r, \tag{35}$$

for some constant α.

B. CGU State Sets With GU Generators

A special class of CGU state sets is *CGU state sets with GU generators* in which the generators $\{\rho_k = \phi_k\phi_k^*, 1 \le k \le r\}$ and the factors ϕ_k are themselves GU. Specifically, $\{\phi_k = V_k\phi\}$ for some generator ϕ, where the matrices $\{V_k, 1 \le k \le r\}$ are unitary, and form a group $\mathcal{Q}$.

Suppose that U_i and V_k commute up to a phase factor for all i and k so that $U_iV_k = V_kU_ie^{j\theta(i,k)}$ where $\theta(i,k)$ is an arbitrary phase function that may depend on the indices i and k. In this case we say that $\mathcal{G}$ and $\mathcal{Q}$ commute up to a phase factor and that the corresponding state set is *CGU with commuting GU generators*. (In the special case in which $\theta = 0$ so that $U_iV_k = V_kU_i$ for all i, k, the resulting state set is GU [24]). Then for all i, k, $\Phi\Phi^*$ commutes with U_iV_k [7], and the SIM factors μ_{ik} are given by

$$\mu_{ik} = \gamma T\phi_{ik} = \gamma TU_iV_k\phi = U_iV_k\bar{\mu}, \tag{36}$$

where $\bar{\mu} = \gamma T\phi$. Thus even though the state set is not in general GU, the SIM factors can be computed using a single generator.

Alternatively, we can express μ_{ik} as $\mu_{ik} = U_i\mu_k$ where the generators μ_k are given by

$$\mu_k = V_k\bar{\mu}. \tag{37}$$

From (37) it follows that the generators μ_k are GU with generating group $\mathcal{Q} = \{V_k, 1 \le k \le r\}$ and generator $\bar{\mu}$. Then for all k,

$$\mu_k^*\phi_k = \bar{\mu}^*V_k^*V_k\phi = \bar{\mu}^*\phi. \tag{38}$$

If in addition,

$$\bar{\mu}^* \phi = \gamma \phi^* T \phi = \alpha I \tag{39}$$

for some α, then combining (34), (38) and (39) with Theorem 1 we conclude that the SIM is optimal. In particular, for a pure-state ensemble, $\bar{\mu}^* \phi$ is a scalar so that (39) is always satisfied. Therefore, for a pure CGU state set with commuting GU generators, the SIM maximizes P_D subject to $P_I = \beta$ for $\beta \geq \beta_{\min}$.

C. Optimal Measurement for Arbitrary CGU States

If the generators ϕ_k do not satisfy (35), or if $\beta < \beta_{\min}$, then the SIM is no longer guaranteed to be optimal. Nonetheless, as we now show, the optimal measurement operators that maximize P_D subject to $P_I = \beta$ are CGU with generating group $\mathcal{G}$. The corresponding generators can be computed very efficiently in polynomial time within any desired accuracy.

Suppose that the optimal measurement operators that maximize

$$J(\{\Pi_{ik}\}) = \sum_{i=1}^{l} \sum_{k=1}^{r} \mathrm{Tr}(\rho_{ik} \Pi_{ik}), \tag{40}$$

subject to

$$P_I(\{\Pi_{ik}\}) = 1 - \frac{1}{lr} \mathrm{Tr} \left(\sum_{i,j=1}^{l} \sum_{k,s=1}^{r} \rho_{ik} \Pi_{js} \right) = \beta, \tag{41}$$

are $\widehat{\Pi}_{ik}$, and let $\widehat{J} = J(\{\widehat{\Pi}_{ik}\})$. Let $r(j,i)$ be the mapping from $\mathcal{I} \times \mathcal{I}$ to $\mathcal{I}$ with $\mathcal{I} = \{1, \ldots, l\}$, defined by $r(j,i) = s$ if $U_j^* U_i = U_s$. Then the measurement operators $\widehat{\Pi}_{ik}^{(j)} = U_j \widehat{\Pi}_{r(j,i)k} U_j^*$ for any $1 \leq j \leq l$ are also optimal. Indeed, since $\widehat{\Pi}_{ik} \geq 0, 1 \leq i \leq l, 1 \leq k \leq r$ and $\sum_{i=1}^{l} \sum_{k=1}^{r} \widehat{\Pi}_{ik} \leq I$, $\widehat{\Pi}_{ik}^{(j)} \geq 0, 1 \leq i \leq l, 1 \leq k \leq r$ and

$$\sum_{i=1}^{l} \sum_{k=1}^{r} \widehat{\Pi}_{ik}^{(j)} = U_j \left(\sum_{i=1}^{l} \sum_{k=1}^{r} \widehat{\Pi}_{ik} \right) U_j^* \leq U_j U_j^* = I. \tag{42}$$

Using the fact that $\rho_{ik} = U_i \rho_k U_i^*$ for some generators ρ_k,

$$\begin{aligned} J(\{\widehat{\Pi}_{ik}^{(j)}\}) &= \sum_{i=1}^{l} \sum_{k=1}^{r} \mathrm{Tr}(\rho_k U_i^* U_j \widehat{\Pi}_{r(j,i)k} U_j^* U_i) \\ &= \sum_{s=1}^{l} \sum_{k=1}^{r} \mathrm{Tr}(\rho_k U_s^* \widehat{\Pi}_{sk} U_s) \\ &= \sum_{i=1}^{l} \sum_{k=1}^{r} \mathrm{Tr}(\rho_{ik} \widehat{\Pi}_{ik}) \\ &= \widehat{J}. \end{aligned} \tag{43}$$

Finally,

$$\begin{aligned} &\mathrm{Tr} \left(\sum_{i,s=1}^{l} \sum_{k,t=1}^{r} \rho_{ik} \widehat{\Pi}_{st}^{(j)} \right) = \\ &= \mathrm{Tr} \left(\sum_{i,s=1}^{l} \sum_{k,t=1}^{r} U_j^* U_i \rho_k U_i^* U_j \widehat{\Pi}_{st} \right) \\ &= \mathrm{Tr} \left(\sum_{i,s=1}^{l} \sum_{k,t=1}^{r} U_i \rho_k U_i^* \widehat{\Pi}_{st} \right) \\ &= \mathrm{Tr} \left(\sum_{i,s=1}^{l} \sum_{k,t=1}^{r} \rho_{ik} \widehat{\Pi}_{st} \right), \end{aligned} \tag{44}$$

so that from (41), $P_I(\{\widehat{\Pi}_{ik}^{(j)}\}) = P_I(\{\widehat{\Pi}_{ik}\})$.

Since the measurement operators $\widehat{\Pi}_{ik}^{(j)}$ are optimal for any j, it follows immediately that the measurement operators $\{\overline{\Pi}_{ik} = (1/l) \sum_{j=1}^{l} \widehat{\Pi}_{ik}^{(j)}, 1 \leq i \leq l, 1 \leq k \leq r\}$ and $\overline{\Pi}_0 = I - \sum_{i,k} \overline{\Pi}_{ik}$ are also optimal. Now, for any $1 \leq i \leq l, 1 \leq k \leq r$,

$$\begin{aligned} \overline{\Pi}_{ik} &= \frac{1}{l} \sum_{j=1}^{l} U_j \widehat{\Pi}_{r(j,i)k} U_j^* \\ &= \frac{1}{l} \sum_{s=1}^{l} U_i U_s^* \widehat{\Pi}_{sk} U_s U_i^* \\ &= U_i \left(\frac{1}{l} \sum_{s=1}^{l} U_s^* \widehat{\Pi}_{sk} U_s \right) U_i^* \\ &= U_i \widehat{\Pi}_k U_i^*, \end{aligned} \tag{45}$$

where $\widehat{\Pi}_k = (1/l) \sum_{s=1}^{l} U_s^* \widehat{\Pi}_{sk} U_s$.

We therefore conclude that the optimal measurement operators can always be chosen to be CGU with the same generating group $\mathcal{G}$ as the original state set. Thus, to find the optimal measurement operators all we need is to find the optimal generators $\{\widehat{\Pi}_k, 1 \leq k \leq r\}$. The remaining operators are obtained by applying the group $\mathcal{G}$ to each of the generators.

Since the optimal measurement operators satisfy $\Pi_{ik} = U_i \Pi_k U_i^*$ and $\rho_{ik} = U_i \rho_k U_i^*$, $\mathrm{Tr}(\rho_{ik} \Pi_{ik}) = \mathrm{Tr}(\rho_k \Pi_k)$, so that the problem (2) reduces to the maximization problem

$$\max_{\Pi_k \in \mathcal{B}} \sum_{k=1}^{r} \mathrm{Tr}(\rho_k \Pi_k), \tag{46}$$

subject to the constraints

$$\begin{aligned} &\Pi_k \geq, \quad 1 \leq k \leq r; \\ &\sum_{i=1}^{l} \sum_{k=1}^{r} U_i \Pi_k U_i^* \leq I; \\ &1 - \frac{1}{r} \mathrm{Tr} \left(\sum_{i=1}^{l} \sum_{k,l=1}^{r} U_i \rho_k U_i \Pi_l \right) = \beta. \end{aligned} \tag{47}$$

Since this problem is a (convex) semidefinite programming problem, the optimal generators Π_k can be computed very efficiently in polynomial time within any desired accuracy [26, 28, 29], for example using the LMI toolbox on Matlab. Note that the problem of (46) and (47) has rn^2 real unknowns and $r+2$ constraints, in contrast with the original maximization (2) subject to (1) and (4) which has lrn^2 real unknowns and $lr+2$ constraints.

We summarize our results regarding CGU state sets in the following theorem:

Theorem 3 (CGU state sets). *Let $\mathcal{S} = \{\rho_{ik} = U_i\rho_k U_i^*, 1 \le i \le l, 1 \le k \le r\}$ be a compound geometrically uniform (CGU) state set on an n-dimensional Hilbert space generated by a finite group $\mathcal{G}$ of unitary matrices and generators $\{\rho_k = \phi_k\phi_k^*, 1 \le k \le r\}$, and let Φ be the matrix of columns $\phi_{ik} = U_i\phi_k$. Then the scaled inverse measurement (SIM) is given by the measurement operators $\Sigma_{ik} = \mu_{ik}\mu_{ik}^*, 1 \le i \le l, 1 \le k \le r$ and $\Sigma_0 = I - \sum_{i,k}\mu_{ik}\mu_{ik}^*$ with*

$$\mu_{ik} = U_i\mu_k$$

where

$$\mu_k = \gamma(\Phi\Phi^*)^{-1}\phi_k,$$

and $\gamma^2 = (1-\beta)/n$. The SIM has the following properties:

1. *The measurement operators $\Sigma_{ik}, 1 \le i \le l, 1 \le k \le r$ are CGU with generating group $\mathcal{G}$;*
2. *The probability of correctly detecting each of the states ϕ_{ik} for fixed k using the SIM is the same;*
3. *If $\phi_k^*(\Phi\Phi^*)^{-1}\phi_k = \alpha I$ for some α and for $1 \le k \le r$, then the SIM maximizes P_D subject to $P_I = \beta$ with $\beta \ge 1 - n\lambda_n$ where λ_n is the smallest eigenvalue of $(1/lr)\sum_{i,k}\rho_{ik}$.*

If in addition the generators $\{\phi_k = V_k\phi, 1 \le k \le r\}$ are geometrically uniform with $U_iV_k = V_kU_ie^{j\theta(i,k)}$ for all i,k, then

1. *$\mu_{ik} = U_iV_k\bar{\mu}$ where $\bar{\mu} = \gamma(\Phi\Phi^*)^{-1}\phi$ so that the SIM operators are CGU with geometrically uniform generators;*
2. *The probability of correctly detecting each of the states ϕ_{ik} using the SIM is the same;*
3. *If $\phi^*(\Phi\Phi^*)^{-1}\phi = \alpha I$ for some α, then the SIM maximizes P_D subject to $P_I = \beta$ with $\beta \ge 1 - n\lambda_n$. In particular, if $\phi = |\phi\rangle$ is a vector so that the state set is a pure-state ensemble, then the SIM maximizes P_D subject to $P_I = \beta$ with $\beta \ge 1 - n\lambda_n$.*

For arbitrary CGU state sets the optimal measurement operators $\widehat{\Pi}_{ik}, 1 \le i \le l, 1 \le k \le r$ that maximize P_D subject to $P_I = \beta$ for any β are CGU with generating group $\mathcal{G}$ and generators Π_k that maximize $\sum_{k=1}^r Tr(\rho_k\Pi_k)$ subject to $\Pi_k \ge 0, 1 \le k \le r$, $\sum_{i,k}U_i\Pi_kU_i^* \le I$, and $Tr\left(\sum_{i=1}^l\sum_{k,l=1}^r U_i\rho_kU_i\Pi_l\right) = r(1-\beta)$.

VII. CONCLUSION

In this paper we considered the optimal measurement operators that maximize the probability of correct detection given a fixed probability β of an inconclusive result, when distinguishing between a collection of *mixed* quantum states. We first derived a set of necessary and sufficient conditions for optimality by exploiting principles of duality theory in vector space optimization. Using these conditions, we derived a general condition under which the SIM is optimal. We then considered state sets with a broad class of symmetry properties for which the SIM is optimal. Specifically, we showed that for GU state sets and for CGU state sets with generators that satisfy certain constraints and for values of β exceeding a threshold, the SIM is optimal. We also showed that for arbitrary GU and CGU state sets and for arbitrary values of β, the optimal measurement operators have the same symmetries as the original state sets. Therefore, to compute the optimal measurement operators, we need only to compute the corresponding generators. As we showed, the generators can be computed very efficiently in polynomial time within any desired accuracy by solving a semidefinite programming problem.

Acknowledgments

The author wishes to thank Prof. A. Megretski and Prof. G. C. Verghese for valuable discussions that lead to many of the results in this paper.

APPENDIX A: NECESSARY AND SUFFICIENT CONDITIONS FOR OPTIMALITY

Denote by Λ the set of all ordered sets $\Pi = \{\Pi_i\}_{i=0}^m, \Pi_i \in \mathcal{B}$ satisfying (1) and (4) with $\beta < 1$, and define $J(\Pi) = \sum_{i=1}^m p_i \mathrm{Tr}(\rho_i\Pi_i)$. Then our problem is

$$\max_{\Pi\in\Lambda} J(\Pi). \tag{A1}$$

We refer to this problem as the primal problem, and to any $\Pi \in \Lambda$ as a primal feasible point. The optimal value of $J(\Pi)$ is denoted by $\widehat{J}$.

To derive necessary and sufficient conditions for optimality, we now formulate a *dual problem* whose optimal value serves as a certificate for $\widehat{J}$. As described in [3], a general method for deriving a dual problem is to invoke the separating hyperplane theorem [36], which states that two disjoint convex sets[39] can always be separated by a

hyperplane. We will take one convex set to be the point 0, and then carefully construct another convex set that does not contain 0, and that captures the equality constraints in the primal problem and the fact that for any primal feasible point, the value of the primal function is no larger than the optimal value. The dual variables will then emerge from the parameters of the separating hyperplane.

In our problem we have two equality constraints, $\sum_{i=0}^{m}\Pi_i = I$ and $\mathrm{Tr}(\Delta\Pi_0) = \beta$ and we know that $\widehat{J} \geq J(\Pi)$. Our constructed convex set will accordingly consist of matrices of the form $-I+\sum_{i=0}^{m}\Pi_i$ where $\Pi_i \in \mathcal{B}$ and $\Pi_i \geq 0$, scalars of the form $\beta - \mathrm{Tr}(\Delta\Pi_0)$, and scalars of the form $r - J(\Pi)$ where $r > \widehat{J}$. We thus consider the real vector space

$$\mathcal{L} = \mathcal{B}\times\mathcal{R}\times\mathcal{R} = \{(S,x,y): \quad S\in\mathcal{B},\ x,y\in\mathcal{R}\},$$

where $\mathcal{R}$ denotes the reals, with inner product defined by

$$\langle (W,z,t),(S,x,y)\rangle = \mathrm{Tr}(WS) + zx + ty. \tag{A2}$$

We then define the subset Ω of $\mathcal{L}$ as points of the form

$$\Omega = \left(-I+\sum_{i=0}^{m}\Pi_i, \beta - \mathrm{Tr}\left(\Delta\Pi_0\right), r - \sum_{i=1}^{m} p_i \mathrm{Tr}(\Pi_i\rho_i)\right), \tag{A3}$$

where $\Pi_i\in\mathcal{B}, \Pi_i\geq 0, r\in\mathcal{R}$ and $r>\widehat{J}$.

It is easily verified that Ω is convex, and $0\notin\Omega$. Therefore, by the separating hyperplane theorem, there exists a nonzero vector $(Z,a,b)\in\mathcal{L}$ such that $\langle (Z,a,b),(Q,c,d)\rangle \geq 0$ for all $(Q,c,d)\in\Omega$, *i.e.,*

$$\mathrm{Tr}\left(Z\left(-I+\sum_{i=0}^{m}\Pi_i\right)\right) + b\left(\beta - \mathrm{Tr}\left(\Delta\Pi_0\right)\right) + \\ +a\left(r-\sum_{i=1}^{m} p_i\mathrm{Tr}(\Pi_i\rho_i)\right)\geq 0 \tag{A4}$$

for all $\Pi_i\in\mathcal{B}$ and $r\in\mathcal{R}$ such that $\Pi_i\geq 0$, $r>\widehat{J}$.

As we now show, the hyperplane parameters (Z,a,b) have to satisfy certain constraints, which lead to the formulation of the dual problem. Specifically, (A4) with $\Pi_i = 0$, $r\to\widehat{J}$ implies

$$a\widehat{J} \geq \mathrm{Tr}(Z) - b\beta. \tag{A5}$$

Similarly, (A4) with $r=\widehat{J}+1$, $\Pi_j = 0$ for $j\neq i$, $\Pi_i = t|x\rangle\langle x|$ for one value $1\leq i\leq m$ where $|x\rangle\in\mathbb{C}^n$ is fixed and $t\to+\infty$ yields $\langle x|Z - ap_i\rho_i|x\rangle\geq 0$. Since $|x\rangle$ and i are arbitrary, this implies

$$Z \geq ap_i\rho_i, \quad 1\leq i\leq m. \tag{A6}$$

With $r=\widehat{J}+1$, $\Pi_j = 0$ for $j\neq 0$, $\Pi_0 = t|x\rangle\langle x|$ where $|x\rangle\in\mathbb{C}^n$ is fixed and $t\to+\infty$, (A4) yields $\langle x|Z - b\Delta|x\rangle\geq 0$, which implies

$$Z\geq b\Delta. \tag{A7}$$

With $\Pi_i = 0, 0\leq i\leq m$, $r\to+\infty$, (A4) implies $a\geq 0$. If $a=0$, then (A5) yields $\mathrm{Tr}(Z)\leq b\beta < b$ and (A7) yields $\mathrm{Tr}(Z)\geq b$. Therefore we conclude that $a>0$, and define $\widehat{X} = Z/a$, $\hat{\delta} = b/a$. Then (A5) implies that

$$T(\widehat{X},\hat{\delta}) \leq \widehat{J}, \tag{A8}$$

where $T(X,\delta) = \mathrm{Tr}(X) - \delta\beta$, (A6) implies that $\widehat{X}\geq p_i\rho_i$ for $1\leq i\leq m$, and (A7) implies that $\widehat{X}\geq\hat{\delta}\Delta$.

Let Γ be the set of $X\in\mathcal{B}$, $\delta\in\mathcal{R}$ satisfying $X\geq p_i\rho_i, 1\leq i\leq m$ and $X\geq\delta\Delta$. Then for any $X,\delta\in\Gamma$, $\Pi\in\Lambda$, we have

$$T(X,\delta) - J(\Pi) = \mathrm{Tr}\left(\sum_{i=1}^{m}\Pi_i(X-p_i\rho_i)\right) \\ +\mathrm{Tr}\left(\Pi_0(X-\delta\Delta)\right)\geq 0. \tag{A9}$$

Since $\widehat{X}\in\Gamma$, from (A8) and (A9) we conclude that $T(\widehat{X},\hat{\delta}) = \widehat{J}$.

Thus we have proven that the dual problem associated with (A1) is

$$\min_{X\in\mathcal{B},\delta\in\mathcal{R}} tr(X) - \delta\beta, \tag{A10}$$

subject to

$$\begin{aligned} X &\geq p_i\rho_i, \quad 1\leq i\leq m; \\ X &\geq \delta\Delta. \end{aligned} \tag{A11}$$

Furthermore, we have shown that there exists an optimal $\widehat{X},\hat{\delta}\in\Gamma$ and an optimal value $\widehat{T} = T(\widehat{X},\hat{\delta})$ such that $\widehat{T} = \widehat{J}$.

Let $\widehat{\Pi}_i$ denote the optimal measurement operators. Then combining (A9) with $\widehat{T} = \widehat{J}$, we conclude that

$$\begin{aligned} (\widehat{X} - p_i\rho_i)\widehat{\Pi}_i &= 0, \quad 1\leq i\leq m; \\ (\widehat{X} - \hat{\delta}\Delta)\widehat{\Pi}_0 &= 0. \end{aligned} \tag{A12}$$

Once we find the optimal $\widehat{X}$ and $\hat{\delta}$ that minimize the dual problem (A10), the constraints (A12) are necessary and sufficient conditions on the optimal measurement operators $\widehat{\Pi}_i$. We have already seen that these conditions are necessary. To show that they are sufficient, we note that if a set of feasible measurement operators Π_i satisfies (A12), then $\sum_{i=1}^{m}\mathrm{Tr}\left(\Pi_i(\widehat{X}-p_i\rho_i)\right) = 0$ and $\mathrm{Tr}\left((\widehat{X}-\hat{\delta}\Delta)\widehat{\Pi}_0\right) = 0$ so that from (A9), $J(\Pi) = T(\widehat{X},\hat{\delta}) = \widehat{J}$.

APPENDIX B: PROOF OF THEOREM 1

In this appendix we prove Theorem 1. Specifically, we show that for a set of states $\rho_i = \phi_i\phi_i^*$ with prior probabilities p_i, if $(1/\gamma)\mu_i^*\psi_i = \alpha I, 1\leq i\leq m$, where $\mu_i = \gamma(\Psi\Psi^*)^{-1}\psi_i = \gamma\Delta^{-1}\psi_i$ are the SIM factors and

$\psi_i = \sqrt{p_i}\phi_i$, then there exists an Hermitian X and a constant δ such that

$$X \geq \psi_i \psi_i^*, \quad 1 \leq i \leq m; \tag{B1}$$
$$X \geq \delta\Delta; \tag{B2}$$
$$(X - \psi_i\psi_i^*)\mu_i\mu_i^* = 0, \quad 1 \leq i \leq m; \tag{B3}$$
$$(X - \delta\Delta)(I - \gamma^2\Delta^{-1}) = 0. \tag{B4}$$

Let $X = \alpha\Delta$ and $\delta = \alpha$. Then (B2) and (B4) are immediately satisfied. Next, since $\alpha I = \psi_i^*\Delta^{-1}\psi_i = \psi_i^*\Delta^{-1/2}\Delta^{-1/2}\psi_i$, it follows that

$$\alpha I \geq \Delta^{-1/2}\psi_i\psi_i^*\Delta^{-1/2}. \tag{B5}$$

Multiplying both sides of (B5) by $\Delta^{1/2}$ we have

$$\alpha\Delta \geq \psi_i\psi_i^*, \tag{B6}$$

which verifies that the conditions (B1) are satisfied.

Finally,

$$(X - \psi_i\psi_i^*)\mu_i = \alpha\gamma\Delta\Delta^{-1}\psi_i - \alpha\gamma\psi_i = 0, \tag{B7}$$

so that the conditions (B3) are also satisfied.

[1] A. Peres, Found. Phys. **20**, 1441 (1990).
[2] A. Peres, *Quantum Theory: Concepts and Methods* (Boston: Kluwer, 1995).
[3] Y. C. Eldar, A. Megretski, and G. C. Verghese, IEEE Trans. Inform. Theory, to appear; also available at quant-ph/0205178 (2002).
[4] A. S. Holevo, J. Multivar. Anal. **3**, 337 (1973).
[5] H. P. Yuen, R. S. Kennedy, and M. Lax, IEEE Trans. Inform. Theory **IT-21**, 125 (1975).
[6] C. W. Helstrom, *Quantum Detection and Estimation Theory* (New York: Academic Press, 1976).
[7] Y. C. Eldar, A. Megretski, and G. C. Verghese, quant-ph/0211111 (2002).
[8] M. Charbit, C. Bendjaballah, and C. W. Helstrom, IEEE Trans. Inform. Theory **35**, 1131 (1989).
[9] M. Osaki, M. Ban, and O. Hirota, Phys. Rev. A **54**, 1691 (1996).
[10] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, Int. J. Theor. Phys. **36**, 1269 (1997).
[11] Y. C. Eldar and G. D. Forney, Jr., IEEE Trans. Inform. Theory **47**, 858 (2001).
[12] C. W. Helstrom, IEEE Trans. Inform. Theory **28**, 359 (1982).
[13] A. Chefles, Phys. Lett. A **239**, 339 (1998).
[14] Y. C. Eldar, IEEE Trans. Inform Theory, to appear; also available at quant-ph/0206093 (2002).
[15] I. D. Ivanovic, Phys. Lett. A **123**, 257 (1987).
[16] D. Dieks, Phys. Lett. A **126**, 303 (1988).
[17] A. Peres, Phys. Lett. A **128**, 19 (1988).
[18] G. Jaeger and A. Shimony, Phys. Lett. A **197**, 83 (1995).
[19] A. Peres and D. R. Terno, J. Phys. A **31**, 7105 (1998).
[20] A. Chefles and S. M. Barnett, Phys. Lett. A **250**, 223 (1998).
[21] A. Chefles and S. M. Barnett, J. Mod. Opt. **45**, 1295 (1998).
[22] C. W. Zhang, C. F. Li, and G. C. Guo, Phys. Lett. A **261**, 25 (1999).
[23] J. Fiurášek and M. Ježek, quant-ph/0208126 (2002).
[24] Y. C. Eldar and H. Bölcskei, IEEE Trans. Inform. Theory, to appear; also available at math.FA/0108096 (2002).
[25] G. H. Golub and C. F. V. Loan, *Matrix Computations* (Baltimore MD: Johns Hopkins Univ. Press, 1996), 3rd ed.
[26] F. Alizadeh, Ph.D. thesis, University of Minnesota, Minneapolis, MN (1991).
[27] F. Alizadeh, in *Advances in Optimization and Parallel Computing*, edited by P. Pardalos (North-Holland, the Netherlands, 1992).
[28] Y. Nesterov and A. Nemirovski, *Interior-Point Polynomial Algorithms in Convex Programming* (Philadelphia, PE: SIAM, 1994).
[29] L. Vandenberghe and S. Boyd, SIAM Rev. **38**, 40 (1996).
[30] M. Ježek, J. Řeháček, and J. Fiurášek, quant-ph/0201109 (2002).
[31] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, quant-ph/0112007 (2002).
[32] E. M. Rains, IEEE Trans. Inform. Theory **47**, 2921 (2001).
[33] K. Audenaert and B. D. Moor, quant-ph/0109155 (2001).
[34] G. D. Forney, Jr., IEEE Trans. Inform. Theory **37**, 1241 (1991).
[35] M. A. Armstrong, *Groups and Symmetry* (New York: Springer-Verlag, 1988).
[36] D. G. Luenberger, *Optimization by Vector Space Methods* (New York, NY: John Willey & Sons, 1968).
[37] Otherwise we can transform the problem to a problem equivalent to the one considered in this paper by reformulating the problem on the subspace spanned by the eigenvectors of $\{\rho_i, 1 \leq i \leq m\}$.
[38] Interior point methods are iterative algorithms that terminate once a pre-specified accuracy has been reached. A worst-case analysis of interior point methods shows that the effort required to solve a semidefinite program to a given accuracy grows no faster than a polynomial of the problem size. In practice, the algorithms behave much better than predicted by the worst case analysis, and in fact in many cases the number of iterations is almost constant in the size of the problem.
[39] A set C is convex if for any $x, y \in C$, $\alpha x + (1-\alpha)y \in C$ for all $\alpha \in [0, 1]$.